

**CENTRE FOR
COMPUTING AND BIOMETRICS**

**Economies of Scale – A Second
Look at the Data**

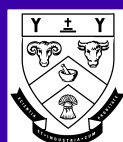
J. Richard Sedcole

Research Report No: 96/05
September 1996

**RESEARCH
REPORT**

ISSN 1173-8405

LINCOLN
UNIVERSITY
Te Whare Wānaka O Aoraki



Centre for Computing and Biometrics

The Centre for Computing and Biometrics (CCB) has both an academic (teaching and research) role and a computer services role. The academic section teaches subjects leading to a Bachelor of Applied Computing degree and a computing major in the BCM degree. In addition it contributes computing, statistics and mathematics subjects to a wide range of other Lincoln University degrees. The CCB is also strongly involved in postgraduate teaching leading to honours, masters and PhD degrees. The department has active research interests in modelling and simulation, applied statistics and statistical consulting, end user computing, computer assisted learning, networking, geometric modelling, visualisation, databases and information sharing.

The Computer Services section provides and supports the computer facilities used throughout Lincoln University for teaching, research and administration. It is also responsible for the telecommunications services of the University.

Research Report Editors

Every paper appearing in this series has undergone editorial review within the Centre for Computing and Biometrics. Current members of the editorial panel are

Dr Alan McKinnon	Dr Keith Unsworth
Dr Bill Rosenberg	Dr Don Kulasiri
Dr Clare Churcher	Mr Kenneth Choo
Dr Jim Young	

The views expressed in this paper are not necessarily the same as those held by members of the editorial panel. The accuracy of the information presented in this paper is the sole responsibility of the authors.

Copyright

Copyright remains with the authors. Unless otherwise stated, permission to copy for research or teaching purposes is granted on the condition that the authors and the series are given due acknowledgement. Reproduction in any form for purposes other than research or teaching is forbidden unless prior written permission has been obtained from the authors.

Correspondence

This paper represents work to date and may not necessarily form the basis for the authors' final conclusions relating to this topic. It is likely, however, that the paper will appear in some form in a journal or in conference proceedings in the near future. The authors would be pleased to receive correspondence in connection with any of the issues raised in this paper. Please contact the authors either by email or by writing to the address below.

Any correspondence concerning the series should be sent to:

The Editor
Centre for Computing and Biometrics
PO Box 84
Lincoln University
Canterbury, NEW ZEALAND
Email: computing@lincoln.ac.nz

Economies of Scale - A Second Look at the Data

J. Richard Sedcole
Centre for Computing and Biometrics, Lincoln University
Canterbury, New Zealand
sedcole@lincoln.ac.nz

The University of Otago Consulting Group were engaged by the Fine Tuning Task Force of the Ministry of Education "...to assist in determining whether the assumed economies of scale in the present school funding formulae for primary and secondary schools indeed exist." A survey of schools was conducted between late 1990 and early 1991 and the results of were published by the Consulting Group in June 1991 (Rothstein, Wilson and MacGibbon, 1991)

The study possessed a number of shortcomings which may mean that the report's conclusions are invalid: the schools were having to adjust to a substantial change in financial management, very few schools were used in the study, the schools were not selected according to a random allocation, data analysis and modelling were inadequate and basic assumptions concerning the validity of marginal cost analysis were overlooked.

1. Introduction

The University of Otago Consulting Group were engaged by the Fine Tuning Task Force of the Ministry of Education "...to assist in determining whether the assumed economies of scale in the present school funding formulae for primary and secondary schools indeed exist." A survey of schools was conducted between late 1990 and early 1991 and the results of this were published by the Consulting Group in June 1991 (Rothstein, Wilson and MacGibbon, 1991).

In so far as the conclusions expressed in the report, as supported by the authors' analysis of their data, are likely to be used by the Ministry of Education in its deliberation over the funding of schools, it is critical that the data have been adequately collected and analysed by the usual tenets of statistical practice.

This critique raises some questions regarding the nature of the data selection, the analysis of the data and the conclusions obtained.

2. The Study

Some details of the data collection and the questionnaires used are given in Rothstein et al (1991): it is sufficient here to note that from a "comprehensive list provided by the Ministry of Education", "representative" schools of the four basic types, primary, intermediate, secondary and area, were chosen in order that a range of sizes occurred within each type. Of the total 66 schools contacted, data from 40 were used in the analysis, but even of these, a number of schools were represented by incomplete data sets. This information is summarised in Table 1:

	Primary	Intermediate	Secondary	Area	Total
Schools contacted	17	15	15	19	66
Schools finally participating	9	10	13	8	40
Schools represented by complete data	6	10	11	3	30

Table 1. Number of different school types contacted and participating in the survey, and number with complete data available.

For each school type the following graphs, based on data collected from the schools, were drawn:

- Surplus per pupil (\$) against school roll
- Non-Government income per pupil (\$) against school roll
- Expenditure per pupil (\$) against school roll
- Percentage fixed expenditure (%) against school roll.

Simple linear and quadratic regressions (apparently) were tried with the data displayed in each of these graphs (with the statistical software "ViewStat"), and where these regressions possessed significant regression coefficients ($p < 0.05$) fitted lines or curves were plotted and in some cases labelled with the coefficients (once wrongly). In most cases no significant relationship was found, and the conclusions pointed to this fact. The results can be summarised in the following table.

	Primary	Intermediate	Secondary	Area
Surplus	-	-	quadratic	-
Non-government income	-	-	-	-
Expenditure	linear	linear	quadratic	-
Percentage fixed	linear	-	linear	-

Table 2. Significant polynomial regressions between (i) Surplus per pupil, (ii) Non-government income per pupil, (iii) Expenditure per pupil, (iv) Percentage of total expenditure that is a fixed cost, against roll size.

It can be seen that linear relationships were found between per pupil expenditure and roll size for primary and intermediate schools, and between percentage fixed expenditure and roll size for primary and secondary schools. Quadratic relationships were found between per pupil surplus and roll size, and between per pupil expenditure and roll size, both for secondary schools. In the second case, this curve was described as "...the typical 'U' shaped average cost curve reflecting economies of scale."

In the case where a quadratic fit occurred for secondary schools per pupil expenditure against school roll, a further graph was drawn, being the marginal cost per pupil against roll size. This seems to be a quadratic curve, but its derivation was not defined. This curve was compared to the formulae then being used by the Ministry of Education in determining its funding to schools which was also plotted on the same graph.

A number of conclusions were listed by the authors of the report of which two are considered here, viz,

- “• There is no significant relationship between per pupil surplus/deficits and school size in primary, intermediate or area schools.
- There is a statistically significant relationship between per pupil surpluses and school size in secondary schools... Schools with either relatively small rolls (under 900 pupils) or relatively large rolls (over 1500 pupils) tend to be more likely to have underspent their Government grant than schools with rolls between 900 and 1500 pupils.”

The report continues with three tables showing the percentage distribution of expenditure for the four school types. Their Table 3.2, for example, cross-classifies the data (expenditure) according to school type (primary, intermediate, secondary and area) and expenditure category (curriculum, administration, and property). For each school type the percentage spent in each category is listed. Their Table 3.1 further subdivides that data to show the percentage of the expenditure within each cell that is due to fixed costs, and what percentage is due to roll related costs.

Their Table 3.3 cross-classifies expenditure by school type and “tagged expenditure category” (maintenance; heat, light and water; minor capital works; vandalism). The median expenditure is given for each cell, as well as the percentage of schools that overspent and underspent in each cell.

3. Criticism of the report

While it is acknowledged that the cost and the time within which the study had to be completed may have precluded a more satisfactory survey from a statistical view point, it is extremely disturbing to see conclusions being drawn where:

- (a) there is no evidence of the data being selected according to a probability (random) criterion,
- (b) there are very few subjects,
- (c) no attempt has been made to assess the variability of the data,
- (d) no standard errors of estimates of polynomial parameters are presented,
- (e) models have been fitted to data without consideration of their appropriateness, and
- (f) basic implied assumptions have not been addressed.

To re-evaluate the analyses presented in the report, the data set was reconstructed by measuring the position of points on the published graphs. Where curves had been fitted, it was possible to test the efficacy of the method. In most cases the results of Rothstein et al (1991) were reproduced, in one instance there was a unit error in the third significant figure. On balance it was felt that residual differences between the reconstructed data set and the original data set of Rothstein et al (1991) would not influence the conclusions of this critique.

We shall deal with each point in turn.

Sampling methodology. The report states that “representative” schools, selected to cover the range of sizes in each type, were chosen, which implies that a “judgemental” or quota sampling method was used rather than any probability (random) sampling. This method can be notoriously untrustworthy, and even with the best of intentions, people who practice selecting “representative” subjects are liable to produce completely spurious conclusions if they try to generalise without looking specifically for negative cases. As it stands, any attempt to make inferences about New Zealand schools’ expenditure patterns from Rothstein et al (1991) can only be considered dubious.

Sample size. The distressingly few subjects are commented on, almost in passing, in the final concluding sentences of the report: a plea that in future, “...a larger and more consistent sample can be used”, and that, “ [The] small number and narrow size range [of area schools] make comparisons within a sample difficult.” The surplus per pupil for area schools is based on just five points, while the best the authors had were the thirteen points used to calculate statistics for secondary schools.

Not mentioned in the report is that there are a little over 1000 state primary schools, about 150 intermediate, about 250 secondary and about 35 state area schools in New Zealand. A few intermediate schools are associated with secondary schools. This list ignores integrated schools. (New Zealand Department of Statistics, 1993).

While sample size can be a much-debated question, and the most efficient sample size cannot be estimated without considering the cost of each subject (Cochran, 1977), we would suggest numbers close to 30 as being a minimum. We shall consider sample size with respect to other points.

Data variability and standard errors. The complete absence of an adequate statistical description of the data set, particularly in describing the variability of the data, could lead the unprepared reader to the wrong conclusion. None of the regression coefficients, or the percentages, were given their appropriate standard errors.

In so far as the authors’ brief (we surmise) was to examine the relationship between various statistics and roll size, the calculation of confidence intervals for mean surplus per pupil, mean non-government income per pupil, etc., where no significant relationship with roll size can be established, may have seemed uncalled for. Nevertheless, confidence intervals for mean responses about regression lines whose slopes are not significantly different from zero can be calculated and, under these circumstances, are valid, emphasising a valuable interpretation of the data.

None of the cross classificatory tables have any standard errors. Although these can be calculated for their Table 3.1, more information is necessary for the other tables. This is especially important for their Table 3.3 since, as displayed, the reader has no idea if, for example, the primary schools that underspent on maintenance underspent by a large or a small margin. Further problems exist with Table 3.3; one is that while the percentages are rounded to the nearest unit, they are displayed to 1 decimal place. For example, the 85.7% of primary schools (six of seven) that underspent on maintenance is listed as “86.0%” In other places in

the table obvious errors exist which purport to show totals of 163%, and that 10.5% of the five area schools underspent on heat, light and water.

Model appropriateness. In some of the cases where data showed a significant relationship between the statistic and roll size, claims were advanced for a quadratic response (Table 2). It seems that Rothstein et al (1991) used the features of "ViewStat" to test polynomial regressions, and simply recorded the highest degree polynomial for which significant coefficients were flagged.

We shall consider some of the graphs displayed in the report and present an alternative model, and hence an alternative conclusion.

Surplus curves.

Figure 1 presents the data for per pupil surplus against roll size for primary schools. No clear relationship exists, but nevertheless a mean of \$75.00 per pupil with a 95% confidence interval of (\$5.80, \$144.20) (which assumes a random sample!) displays a clear indication of the variability of the data.

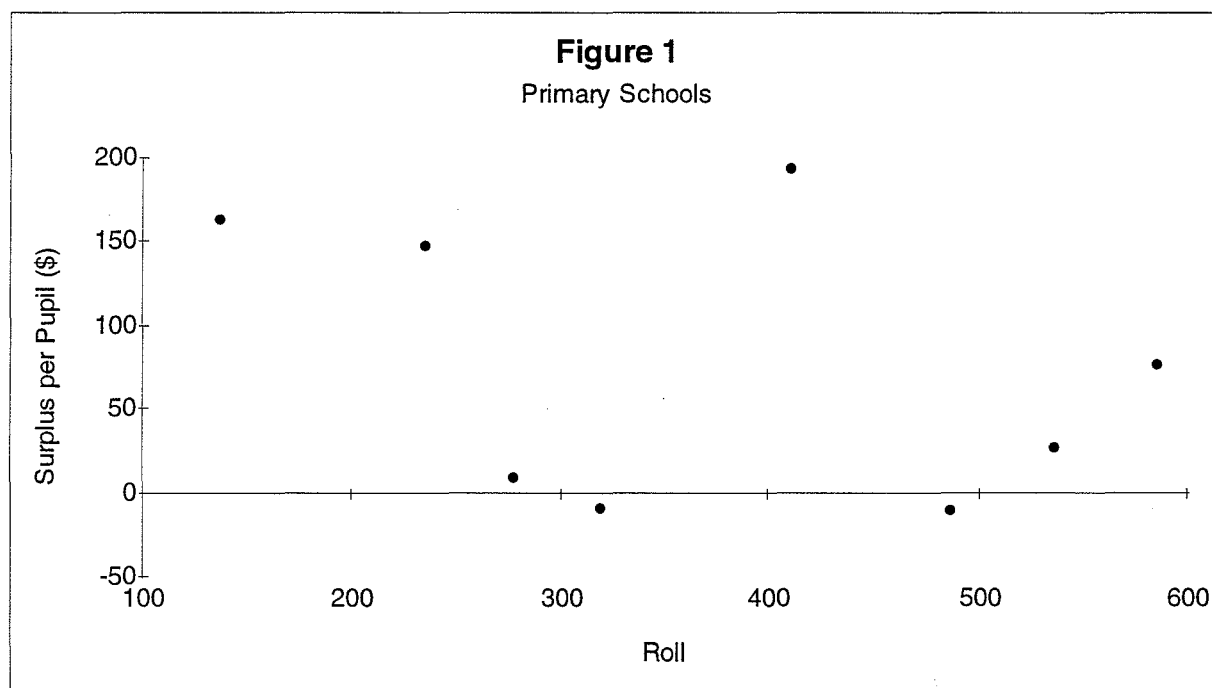
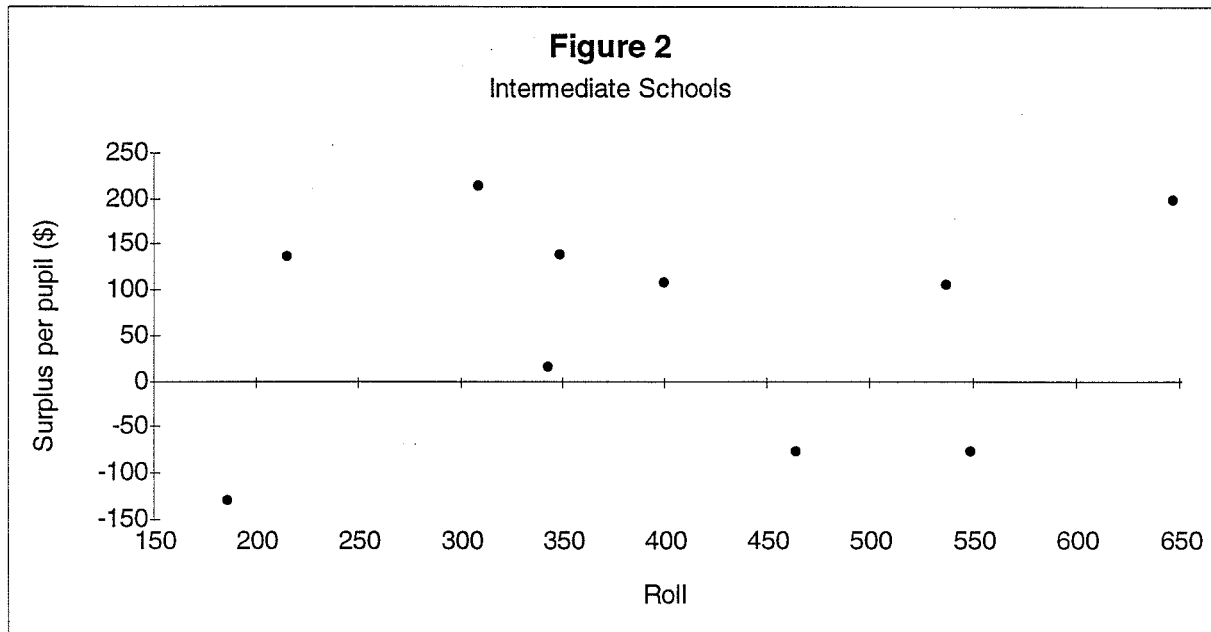
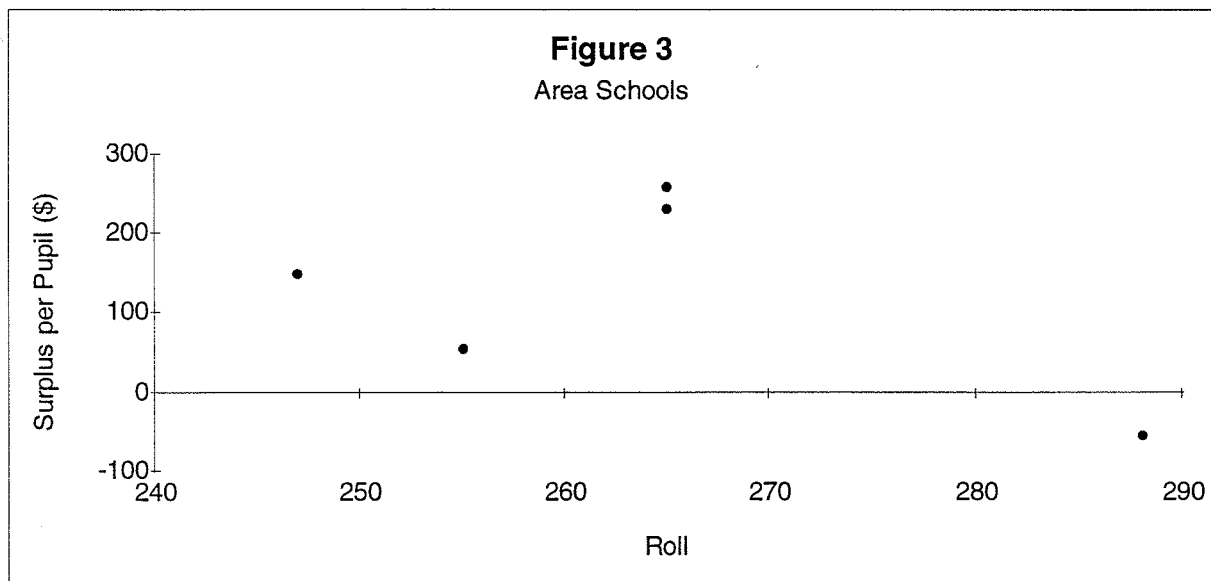


Figure 2 depicts the same statistics for intermediate schools. Again, there is clear scatter and no indication of a linear trend.



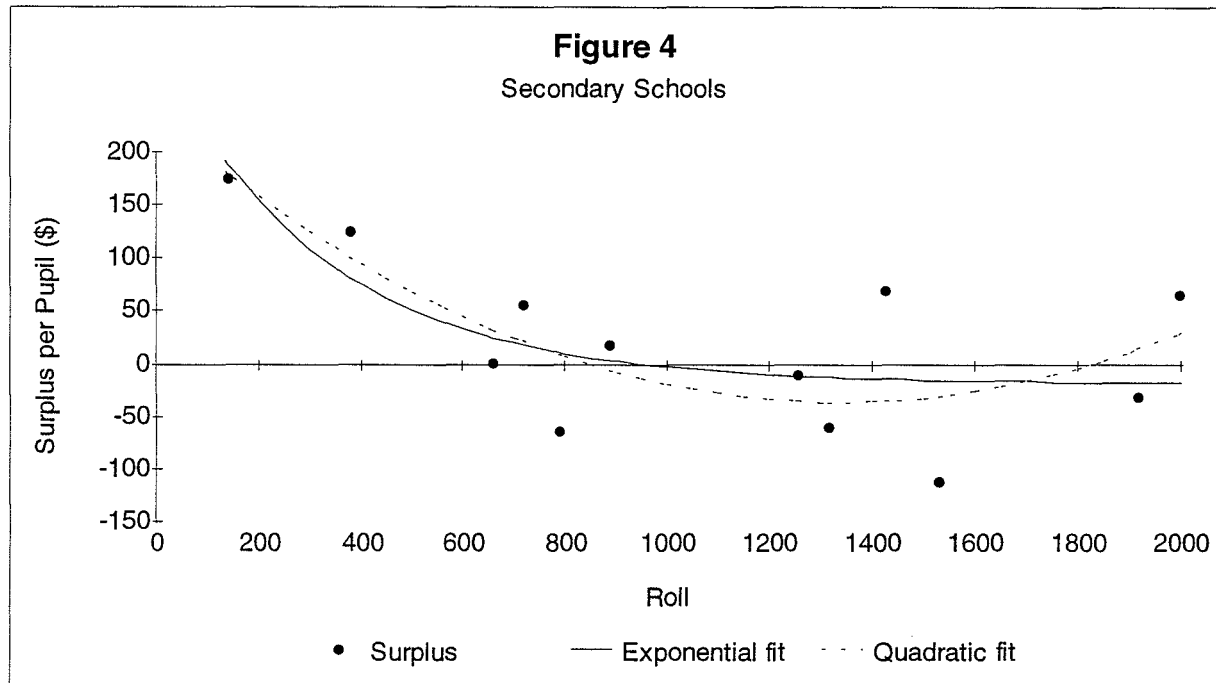
The case for area schools (Figure 3), however, suggests that large schools tend to have a lower per pupil surplus, but the marked scatter and very few points result in a non-significant regression.



In Figure 4, where the per pupil surplus is plotted against roll size for secondary schools, this relationship was modelled by Rothstein et al (1991) by a quadratic expression, although no coefficients for the presumed polynomial were given. (We estimate the quadratic to be

$$y = 231.4 [\pm 60.9] - 0.401r [\pm 0.123] + 0.000150r^2 [\pm 0.0000542], \text{ where } r \text{ is the roll size}).$$

Their claim that these data show a quadratic response is being questioned here. If the two points for the smallest schools are removed from the data, no linear or quadratic response remains and the upward swing implied by the curve cannot be substantiated by the remaining data.



A more appropriate model may be one that decreases towards a fixed asymptote as roll increases - such as the exponential decay curve (an old favourite of biometricians). Non-linear model fitting does not seem to be part of the ViewStat suite, and a lack of experience would not have lead the authors to think of the exponential curve. Fitting the model

$$y = a + be^{-cr}$$

where y is the per pupil surplus (\$), r is the roll size and a, b, c estimated parameters, leads to the result shown in Figure 4. Such calculations are generally not part of the more elementary statistical software. A couple of macros were used with MINITAB to fit the exponential curve and calculate the parameter estimates and their standard errors which are:

$$a = -18.7 (\pm 31.3), b = 316 (\pm 137) \text{ and } c = -0.00302 (\pm 0.00210).$$

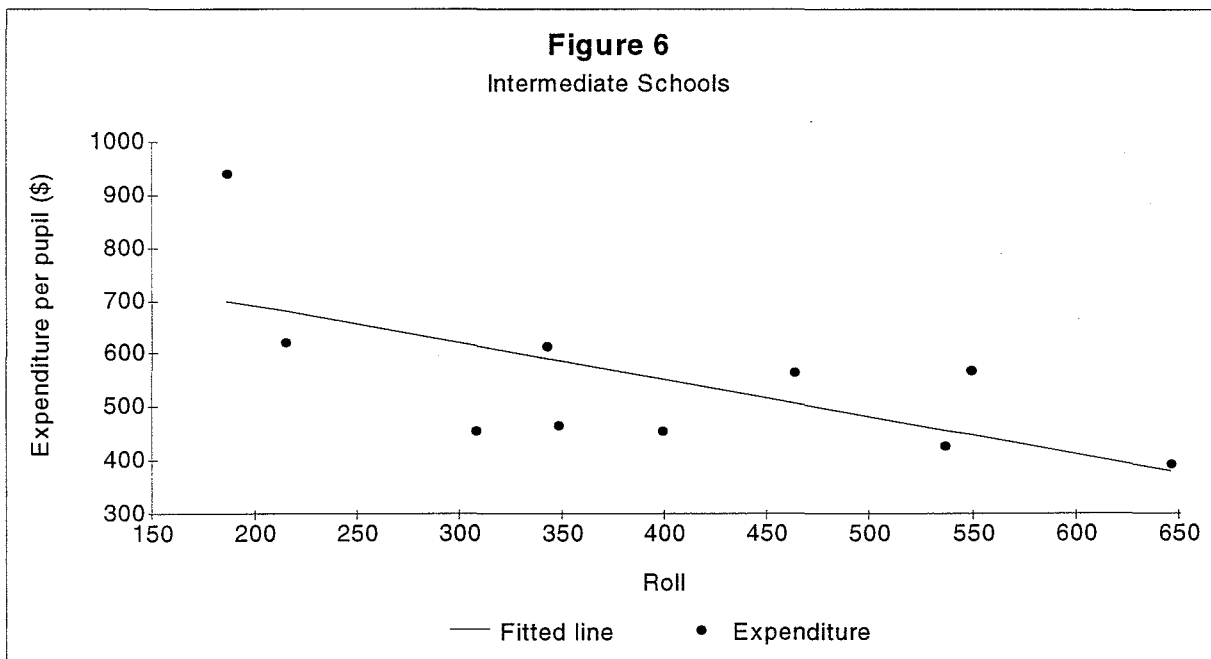
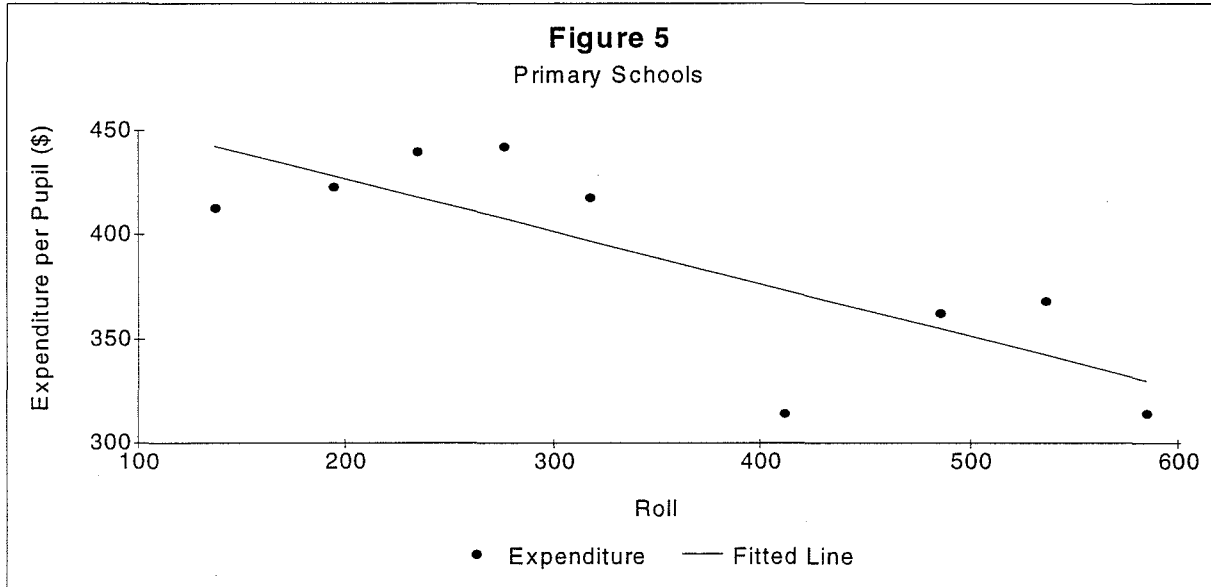
The quadratic and exponential curves are displayed in Figure 4.

In view of the wide scatter of points and the fact that both curves are influenced by the two first points, neither curve could be considered a satisfactory model of the data. The exponential curve, however, avoids the perhaps false implication that larger schools have increasing surpluses per pupil.

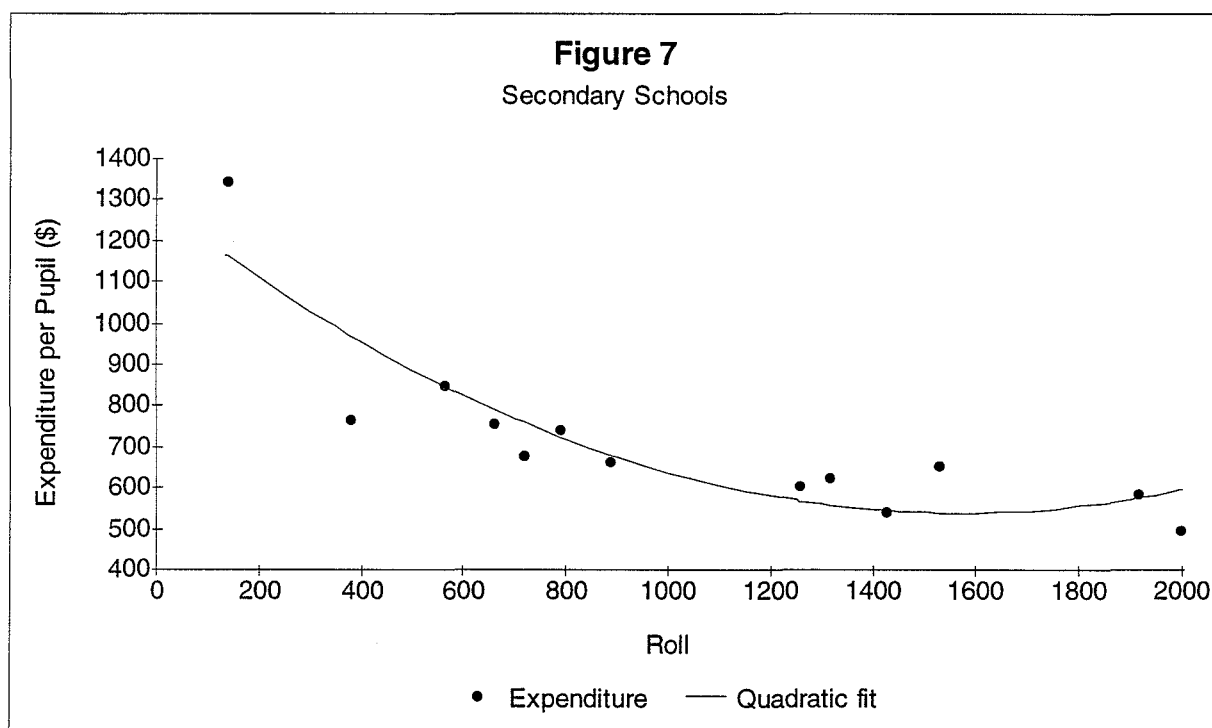
Average and marginal cost curves.

Average cost per pupil (i.e., expenditure per pupil) and marginal cost per pupil models form the central discussion in Rothstein et al (1991). Obviously, these factors are of prime interest to the Ministry of Education. It is equally obvious for the sake of the schools' financial health that the true relationship between roll size and average cost or marginal cost per pupil is not completely lost owing to the application of an inappropriate or naive model.

The relationship between per pupil expenditure and roll size for primary and intermediate schools was modelled by simple linear regressions (Figures 5 and 6).



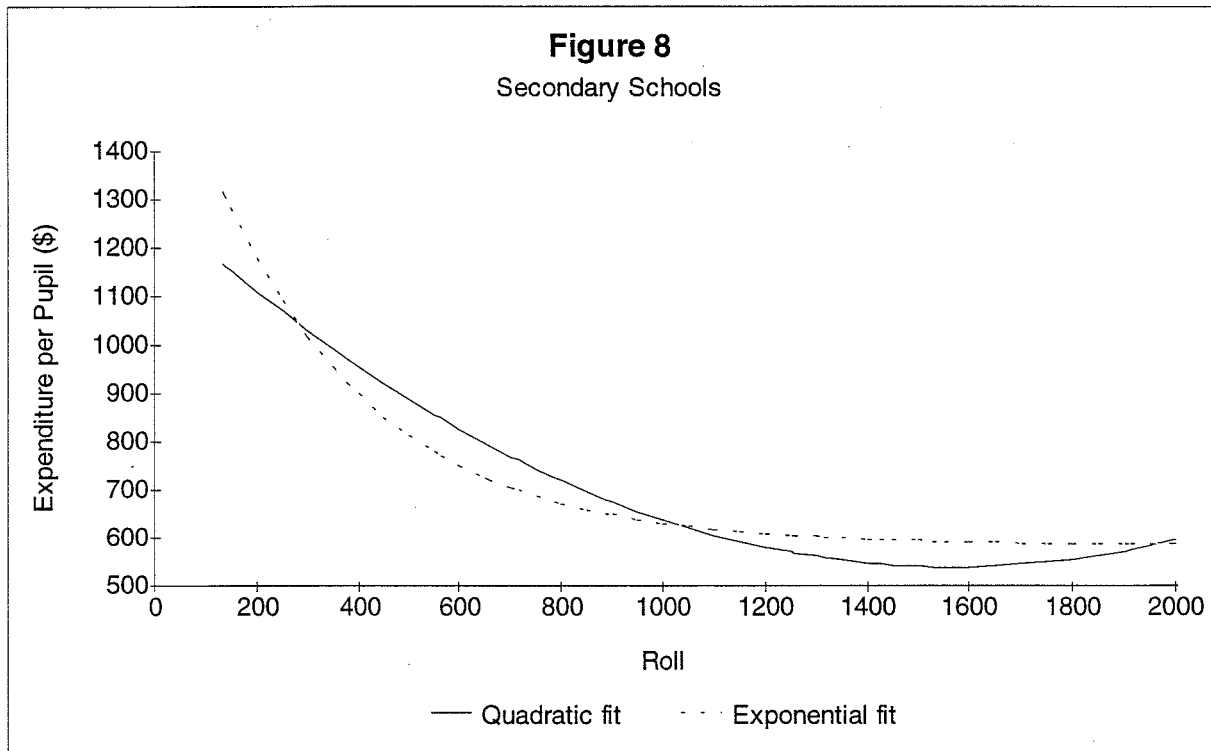
In both cases, visual inspection of the data should have alerted the authors of the report that the data is unsatisfactorily represented by simple linear regressions in spite of the significance of the coefficients. In the case of intermediate schools the value for the smallest school is a clear outlier. Statistical software of more sophistication than ViewStat (for example, MINITAB) possess data checking routines that would identify this point as lying a significant deviation from the fitted regression line. This is a valid datum point, we surmise, and its removal from the data set (which, incidentally, results in a non-significant regression) or ignoring its deviation, are both unsatisfactory approaches to modelling the relationships between per pupil expenditure and roll size.



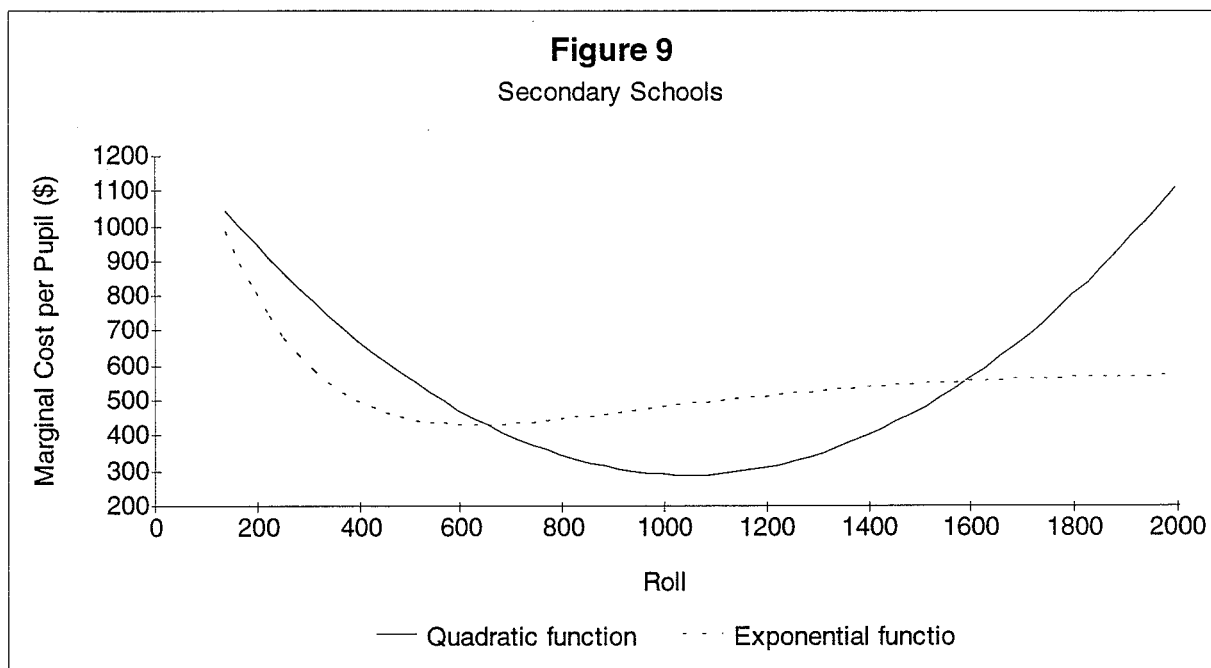
In Figure 7 the authors display the data for per pupil expenditure and roll size for secondary schools, and a quadratic curve fitted to the data. However, like the per pupil surplus graph, one point governs the existence of the quadratic curve. There is again no evidence of an upswing in the data for larger schools. Rothstein et al state that their literature review showed that "... the cost curve was shown to be usually 'U' shaped..." but their reference to Fox (1981) did not give any other information at all. If the data of Fox refers to schools with rolls ranging over quite different values from those analysed here, the implication is obvious. As it stands, we vigorously dispute conclusion of Rothstein et al (page 30) that "per pupil expenditure [for secondary schools] followed the typical 'U' shaped average cost curve reflecting economies of scale."

Marginal cost is nowhere defined or referenced by Rothstein et al (1991), but the shape of their curve is consistent with a usual definition of marginal cost, being the differential (with respect to r) of the function $r f(r)$ where r is the school roll and $f(r)$ is the fitted quadratic curve to per pupil expenditure by roll size. This expression, however, is highly sensitive to the type of curve, $f(r)$, and quadratic versions of $f(r)$ give quite different pictures of marginal costs

(especially for values of minimum marginal cost) than other curvilinear versions. The exponential curve, with $a = 584.5$ (± 11.5), $b = 1127.3$ (± 54.2) and $c = -0.003197$ (± 0.000226), is shown alongside the quadratic curve in Figure 8.



The marginal cost curves associated with these two functions is displayed in Figure 9.



A well fitting model ($R^2 = 0.92$) is of the form:

$$y = a + b/r$$

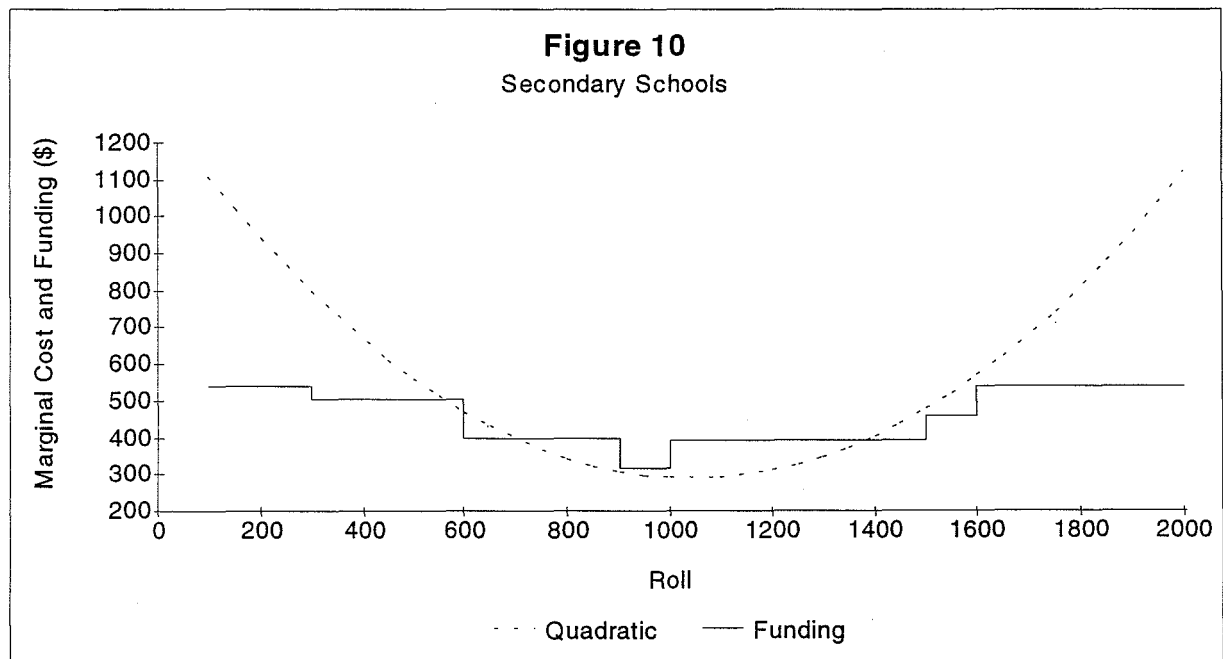
and here $a = 533$ (± 23), and $b = 114063$ (± 9885), and the marginal cost curve derived from this is a *constant* \$533 across all roll sizes. A less well fitting curve ($R^2 = 0.86$) regresses the expenditure on the logarithm of roll size, leading to a marginal cost curve that is parallel to the expenditure curve. These various models emphasize the crucial point that far too few schools feature in this report. There is *a priori* no reason to select one curve over another, even amongst a set of curves with the same number of parameters, but the conclusion about marginal cost - and schools' funding - is a function of the type of model chosen! Until the average cost curve can be well characterised empirically by good data (and the definition of "good" could be an extended paper - perhaps even a conference - in itself), the fitting of any model needs to be heavily qualified by caveats alerting the reader to its arbitrary selection.

As an aside, we tried to calculate the marginal costs directly from the data:

$$m_i = (c_{i+1} r_{i+1} - c_i r_i) / (r_{i+1} - r_i)$$

where c_i is the expenditure per pupil, and r_i is the roll size of the i -th school, ordered so that $r_{i+1} > r_i$. The complete absence of an identifiable relationship would stay the hand of any statistician, and serves to emphasize the reservation about the arbitrariness of the choice of marginal cost functions.

Basic assumptions. There remains to be discussed, however, the whole validity of the entire concept. As almost an aside, the authors comment that "... in general, schools will spend to the extent of their income...". The so-called 'U-shaped' cost curve does not reflect economies of scale - it merely reflects the funding. Figure 10 shows the quadratic marginal cost curve for secondary schools and the Ministry of education funding formula.



In so far as schools' income is based on the Ministry of Education funding formula, it would be difficult to determine on expenditure pattern alone if a school is more effectively teaching 400 or 1400 students. Perhaps the real measure should be the effectiveness of learning (a variable we prefer not to attempt to define here.) Nevertheless, an implied assumption, absolutely critical to the interpretation of the analysis, is that the pupils are equally educated, regardless of the amount of money a school receives.

If a clear pattern of non-government income (NGI) existed, then certain sized schools may show up as being more effective at fund raising than others. No relationship between NGI and roll size is discernible and we are left with the relationship:

Expenditure per pupil = Ministry funding per pupil + random variable where the random variable is estimated from the data to have:

School	mean(\$)	standard deviation (\$)
Primary	78	27
Intermediate	161	92
Area	123	18
Secondary	153	102

and an obvious lower bound of zero (in the absence of deficit provisions).

Just to emphasize the critical issue that Rothstein et al (1991) seem to have missed:

The Ministry of Education funding formula that was in place at the time the survey was taken already has an implied least cost concept built in. Schools spend what they get ! Hence, any analysis of the data will show this as a quadratic-like response, unless the non-government income is a clear function of size. This is NOT the case.

4. Conclusion

In conclusion, a number of disconcerting points are stated.

Firstly, we must emphatically concur with the authors that the data were collected during a period of great change in the financial management of schools, and the variation induced by various schools in their approach to an unknown regime hides, and indeed very likely distorts, any clear pattern of economies of scale.

Secondly, the very few schools selected makes conclusions quite untrustworthy, and considering the potential amount of money being dispersed via the Ministry of Education's funding formula, more resources should have been made available to ensure an adequate sample was selected.

Thirdly, the failure to use appropriate sampling methodology invalidates any attempt to make inferences about school funding generally.

Fourthly, the analysis of the data could most charitably be described as naive. The use of software lacking adequate data diagnostic features is only partly to blame; a failure to visually inspect the data and assess heuristically the adequacy and implications of polynomials was a substantial contribution to the problem.

Fifthly, the failure to describe the variability of the data, by the appropriate standard deviations, standard errors or confidence intervals, give the reader no clue about the variation. If the authors felt that the intended readers were likely to be statistically naive, a few appropriate explanations would have at least warned the readers of the variability associated with the data.

Sixthly, the marginal cost curve is, in fact, no such thing, being merely a curve fitted to the funding formula plus noise (i.e., external fundraising). Even if we allow the marginal cost curve validity, the nature of the curve is heavily dependent on the nature of the average cost curve, and the choice of the inappropriate model, due to variability or inadequacy of data, or naivety of analysis, has considerable consequence to schools funded according to the wrong protocol. This is particularly concerning because, in the present economic direction taken by the Government, increased overall funding to schools is unlikely - rather the distribution of the current funding will be altered, based on deliberations by the Fine Tuning Task Force. If their conclusions are based on the University of Otago Consulting Group Report, then clearly secondary schools with rolls about the 600 pupil level will be disadvantaged relative to secondary schools with rolls about 1100.

In view of the absence of any acknowledgment in the report of statistical assistance, we are left to conclude that none was sought: if the Consulting Group is stationed actually on the Otago University Campus, the Group had within five minutes' walk access to statistical advice equal to anywhere in the world.

5. References

Cochran, W.G. 1977: *Sampling Techniques* (3rd ed.). New York. John Wiley and Sons, Inc.

Fox, W.F. 1981: Reviewing Economies of Size in Education. *Journal of Education Finance* 6, 273 - 296.

New Zealand Department of Statistics 1993: *New Zealand Official Yearbook, 1993*, Government Printer, Wellington, New Zealand.

Rothstein, D.S.; Wilson, A.C.; MacGibbon, L.A.C. 1991: *Economies of Scale in the Funding Formulae of Primary and Secondary Schools - National Study*, University Consulting Group, University of Otago.